Provincial Report

Grade 3 English Language Arts Achievement Test

October 1985

Student Evaluation

Alberta EDUCATION

DISTRIBUTION:

Superintendents of Schools
School Principals and Teachers
The Alberta Teachers' Association
Alberta School Trustees' Association
Alberta Education
General Public Upon Request

EXECUTIVE SUMMARY

Description of the Test

The Grade 3 English Language Arts Achievement test was a two-part test. Part A: Expressive Language - Writing consisted of an assignment in which students were required to write a story based on a story starter. Students were given one hour for writing Part A. Part B: Receptive Language - Reading consisted of 36 multiple-choice questions linked to brief reading selections. Included in Part B were six questions from the Minister's Advisory Committee on Student Achievement (MACOSA) Test, 1978. Students were given 45 minutes to complete Part B.

Administration

The test was administered on June 10, 1985 to 31 232 students enrolled in Grade 3 English Language Arts. Part A: Expressive Language - Writing was scored by 126 Grade 3 teachers under the supervision of Student Evaluation Branch personnel. The students' stories were marked according to five scales: Content, Development, Sentence Structure, Vocabulary, and Conventions. Part B: Receptive Language - Reading was machine scored.

Results

The provincial average score for combined Parts A and B (each worth 50%) was 66.9% and the standard deviation[1] was 14.4. Overall, 89% of the students achieved above the minimum standard set for this test[2].

Part A: Expressive Language - Writing: Overall achievement on the writing assignment was considered acceptable, with no fewer than 71.3% of the students scoring Satisfactory or better on any one marking scale. The highest achievement was on Vocabulary, with 79.5% of the students achieving a Satisfactory or better score; the lowest was on Development, with 71.3% of the students achieving a Satisfactory or better score. Teachers were pleased with the quality and creativity of the students' writing. The students wrote well in the narrative form, and usually enhanced their stories with precise vocabulary, and creative and interesting details.

Overall, 86% of the students achieved above the minimum standard set for Part A of the test.

The overall provincial average score for the written response assignment was 62.7% and the standard deviation was 4.2.

Part B: Receptive Language - Reading: The provincial average score for the 36 multiple-choice questions was 71.0% and the standard deviation was 5.9. Eighty-nine per cent of the students achieved above the minimum standard set for Part B of the test.

A greater percentage of the 1985 Grade 3 students correctly answered the six MACOSA questions than did their 1978 counterparts. The average score was 74.7% in 1985 as compared to the average 1978 score of 69.7%.

---

[1]A measure of variability of scores. In a normal distribution 68% of the students' scores would fall within one standard deviation of the average.
[2]Minumum, acceptable level of achievement; see standard setting p. 7.

86-05558/

TABLE OF CONTENTS

	PAGE
LIST OF TABLES	i
ACKNOWLEDGMENTS	ii
CHAPTER 1: THE ACHIEVEMENT TESTING PROGRAM	1
Exemptions from the Achievement Testing Program	1
CHAPTER 2: TEST DESIGN, DEVELOPMENT, AND DESCRIPTION	2
Test Design and Development	2
Test Description	3
Part A: Expressive Language - Writing	3
Reporting Categories for Part A	3
Part B: Receptive Language - Reading	5
Reporting Categories for Part B	5
Cognitive Levels for Part B	5
CHAPTER 3: ADMINISTRATION OF THE TEST	7
Determination of the Student Population	7
Administration	7
Data Collection	7
Standard-Setting for the Grade 3 English Language Arts Test	7
CHAPTER 4: SCORING OF PART A: EXPRESSIVE LANGUAGE - WRITING	9
Organization of Markers	9
Training	9
Scoring	9
Reliability Reviews	10
CHAPTER 5: RESULTS AND OBSERVATIONS	11
Test Results	11
Results for Part A: Expressive Language - Writing	13
Results for Part B: Receptive Language - Reading	15
Discussion of Selected Questions	20
Summary of Observations	24
CHAPTER 6: GUIDE TO THE INTERPRETATION OF JURISDICTION RESULTS	25
Differences Between Jurisdiction and Provincial Averages	25
Absentee Rates	28
APPENDIX A: GRADE 3 SCORING GUIDES: EXPRESSIVE LANGUAGE - WRITING	29
APPENDIX B: GRADE 3 SAMPLE SCORE SHEET	35

LIST OF TABLES

TABLE		PAGE
1	Grade 3 English Language Arts Achievement Test Part A: Expressive Language - Writing Blueprint	4
2	Distribution of Reading Selection Types	5
3	Grade 3 English Language Arts Achievement Test Part B: Receptive Language - Reading Blueprint	6
4	Grade 3 Language Arts Achievement Test Frequency Distribution of Total Scores	11
5	Part A: Expressive Language - Writing Percentage Distribution of Scores By Reporting Category	13
6	Part A: Expressive Language - Writing Frequency Distribution of Scores	14
7	Results for Part B: Receptive Language - Reading By Reporting Category	15
8	Results for Part B: Receptive Language - Reading By Cognitive Level	16
9	Part B: Receptive Language - Reading Frequency Distribution of Scores	17
10	Comparison of Performance on MACOSA Questions Part B: Receptive Language - Reading	18
11	Part B: Receptive Language - Reading Question Response Frequencies	19
12	Distribution of Jurisdiction Levels of Achievement on the Grade 3 English Language Arts Achievement Test	27

ACKNOWLEDGMENTS

The successful administration of the Grade 3 English Language Arts Achievement Test would have been impossible without substantial contributions from many people. The cooperation of administrators, teachers and students is greatly appreciated.

The advice received from the Test Review Committee regarding design, development, and reporting has been particularly valuable in the implementation of the Achievement Testing Program. This Committee has representation from:

The Alberta Teachers' Association
The Conference of Alberta School Superintendents
The Universities
Alberta Education

The contribution made by this group is gratefully acknowledged.

The technical expertise provided by Dr. T. O. Maguire, Professor, Division of Educational Research Services, University of Alberta, has also contributed greatly to the advancement of the Achievement Testing Program, and his work in this area is acknowledged and appreciated.

George H. Bevan
Director
Student Evaluation Branch

Chapter 1

THE ACHIEVEMENT TESTING PROGRAM

The Achievement Testing Program provides Alberta Education, school jurisdictions, and the public with information, significant at the provincial and local levels, about student knowledge, understanding, and skills in relation to program objectives.

The achievement tests are specific to the program of studies prescribed by the Minister of Education. Curriculum specifications for each subject area, provided by the Curriculum Branch and the Language Services Branch of Alberta Education, identify the major content areas, the specific learning objectives within each area, and the emphasis that each objective is to receive. The test questions reflect these curriculum specifications.

The achievement tests are administered on a cyclical basis in four subject areas: language arts, social studies, mathematics, and science, and at three grade levels: 3, 6, and 9. In 1985, achievement tests were administered in Grade 3 English Language Arts, Grade 6 Social Studies, and Grade 9 Science.

Following the achievement test administration in June of each year, the results are reported to each school jurisdiction. These district profiles include results for each school and each student, but individual statements of results are not issued to students.

This report is designed to assist school jurisdictions in interpreting their achievement test results.

Exemptions from the Achievement Testing Program

Under normal circumstances, the following students are exempt from achievement testing:

- Students participating in Special Education programs
- Students in classes in which the subject being tested has been cycled and taught in an alternate year
- Students in classes in which the subject being tested has been taught in an alternate semester
- Students enrolled in English as a Second Language programs

Chapter 2

TEST DESIGN, DEVELOPMENT, AND DESCRIPTION

Test Design and Development

There were a number of stages in the development of the Grade 3 English Language Arts Achievement Test: preparation of curriculum specifications; development of test design and questions; field testing and revision of questions; construction and administration of a pilot test; and preparation of the final test. After key phases of development, the test was reviewed by a Test Review Committee.

The Curriculum Branch of Alberta Education prepared curriculum specifications that identified the major content areas, the specific objectives within each area, and the emphasis each is to receive in the classroom. The curriculum specifications were distributed to all school jurisdictions in the province in the publication *Grade 3 Language Arts Curriculum Specifications* (1982). From the Curriculum Specifications, the Student Evaluation Branch of Alberta Education selected those specifications that could best form the basis for a paper-and-pencil test and represent the important curricular emphases.

The Student Evaluation Branch of Alberta Education developed a test blueprint and reporting categories based on the selected curriculum specifications. These were then presented to a Test Review Committee. The test design and blueprints as well as the *Grade 3 Curriculum Specifications* guided the development of test questions.

Test questions were developed by Grade 3 English Language Arts teachers from all parts of the province under the supervision of the Student Evaluation Branch. These questions were field-tested and if necessary, revised. Next, a pilot test was constructed, reviewed by the Test Review Committee, and administered.

The final test was constructed from those field and pilot test questions that best reflected curricular intent and test design.

The test design, blueprints, scoring guides, and sample questions, were distributed to all school jurisdictions in the province in the publication *Student Achievement Testing Program: Grade 3 Language Arts* (Student Evaluation Branch Bulletin Volume 4, Number 10, September 1984).

Test Description

The Grade 3 English Language Arts Achievement Test consisted of two parts: Part A: Expressive Language - Writing, and Part B: Receptive Language - Reading.

Part A: Expressive Language - Writing

Part A required students to write a single narrative assignment based on a story starter. Students were allowed one hour for writing Part A.

Students were not allowed to use a dictionary while writing Part A. Space was provided for listing ideas and for completing the story.

Reporting Categories for Part A

To provide information about student writing that is meaningful, students' responses were evaluated in terms of particular writing components. Components evaluated were: Content - the selecting of details appropriate to purpose (details descriptive of character and setting, and narrative details about actions and events); Development - the ordering of details into a coherent whole; Sentence Structure - the control of sentence type, length, and structure; Vocabulary - the selection of words and expressions; and Conventions - the control of spelling, punctuation, and capitalization.

These components or reporting categories are presented on page 4.

Table 1

Grade 3 English Language Arts Achievement Test

Part A: Expressive Language - Writing

Blueprint

REPORTING CATEGORY (Scoring Guide)	DESCRIPTION OF WRITING ASSIGNMENT	RANGE OF MARKS
CONTENT (Selecting Details Appropriate to Purpose) Events should be plausible and appropriate to the student's purpose for communicating. The student should be able to select appropriate details to describe characters and setting. DEVELOPMENT (Ordering Details into a Coherent Whole) The student should be able to place events in a coherent sequence. SENTENCE STRUCTURE (Structuring Sentences Appropriately) The student should be able to use a variety of sentence structures appropriately. VOCABULARY (Selecting and Using Words and Expressions Appropriately) The student should be able to use words and expressions appropriately. CONVENTIONS (Following the Conventions of Written Language Appropriately) The student should be able to communicate clearly by adhering to appropriate spelling, punctuation, and capitalization.	The writing assignment follows a story starter that is to be read by the students. The assignment sets a specific writing task, but allows students to use imagination to select supporting details and events.	In each reporting category, students receive a score within the following range: 5 - EXCEPTIONAL 4 - PROFICIENT 3 - SATISFACTORY 2 - LIMITED 1 - POOR Ins - INSUFFICIENT

Part B: Receptive Language - Reading

Part B consisted of 30 questions based on six reading selections, and six discrete questions re-administered from the last Grade 3 Language Arts Achievement Test administered by the Minister's Advisory Committee on Student Achievement in 1978. Information on the numbers and types of reading selections in Part B follows in Table 2.

Table 2

Distribution of Reading Selection Types

Reading Selection Type	Number of Selections	Number of Questions
Discrete Questions	-	6
Fiction	4	19
Non-fiction	1	5
Poetry	1	6
Total	6	36

Students were allowed one hour for writing Part B. Use of a dictionary was not permitted.

Reporting Categories for Part B

Questions were grouped into five reporting categories or subtests: Main Idea (7 questions), Supporting Detail (7 questions), Vocabulary (6 questions), Predictions (8 questions), and Conclusions (8 questions). Each reporting category required a minimum of six questions so that reliable statistics could be obtained. The reporting categories used in Part B are given in Table 3, page 6.

Cognitive Levels for Part B

Questions were also classified according to two cognitive levels: Literal Understanding (10 questions), and Inferential Understanding and Evaluation (26 questions). By considering cognitive level when developing a test, the Student Evaluation Branch attempts to ensure that a variety of mental activities will be used by students as they write the test. Questions listed under Literal Understanding are expected to be answered using skills of recall and recognition; those listed under Inferential Understanding and Evaluation are expected to elicit skills of analysis, interpretation, extrapolation, and judgment.

The classification of the questions for each reporting category for each cognitive level is shown in Table 3.

Table 3

Grade 3 English Language Arts Achievement Test

Part B: Receptive Language - Reading

Blueprint

REPORTING CATEGORY	COGNITIVE LEVEL		TOTAL
	LITERAL UNDERSTANDING RECALL AND RECOGNITION	INFERENTIAL UNDERSTANDING, EVALUATION	
1. MAIN IDEA The student should be able to determine the main idea of a reading selection.	(2) 31,34	(5) 5,9,14, 19,30	7
2. SUPPORTING DETAIL The student should be able to understand supporting details found in reading selections and evaluate supporting details in terms of the main idea.	(5) 16,22,27 32,35	(2) 2,6	7
3. VOCABULARY The student should be able to recall the meanings of words and expressions, infer word meaning from context, and evaluate appropriateness of word usage.	(3) 1,20,28	(3) 10,15,21	6
4. PREDICTIONS The student should be able to determine the outcome of a series of events.	--	(8) 4,8,12,18, 24,29,33,36	8
5. CONCLUSIONS The student should be able to draw appropriate conclusions from details and ideas present in reading selections, and evaluate the relative importance of concluding statements.	--	(8) 3,7,11,13, 17,23,25,26	8
TOTAL	10	26	36

Chapter 3

ADMINISTRATION OF THE TEST

Determination of the Student Population

The larger school jurisdictions could elect to test all Grade 3 English Language Arts students or just students from randomly selected schools. No jurisdictions opted for sampling. Therefore, the only students who did not write the Grade 3 English Language Arts Achievement Test were those who were normally exempted (see page one) and any student absent from school on the date of administration.

Administration

Jurisdictions were requested in April to report the number of students enrolled in Grade 3 Language Arts in each school. In May, letters were sent to superintendents and principals regarding the test schedule, proceedures for test administration and requiremnets for returning test materials. Information addressed to the teachers related to the administration of the test and the return of test materials. Jurisdictions were sent the appropriate number of tests and administration instructions, packaged according to school. Immediately after the test was administered, teachers were instructed to collect all test booklets and answer sheets and return them to the principal for forwarding to school board offices. School Boards were responsible for sending the test booklets and answer sheets to the Student Evaluation Branch.

Staff from the Regional Offices of Alberta Education supervised the administration of the test in private schools.

Data Collection

A total of 1102 schools from 143 public and separate school jurisdictions returned scorable booklets for 31 232 students. A total of 88 schools from 84 private jurisdictions returned scorable booklets for 798 students.

Standard-Setting for the Grade 3 English Language Arts Test

While provincial averages are useful for comparing the scores of groups of students in a particular school or jurisdiction with provincial levels of achievement, the average score does not necessarily indicate whether the students in the province did as well as they should. In other words, a test score by itself has limited meaning without comparison to a standard. Tests vary in difficulty: a raw score of 18/36, for example, could represent very high achievement on one test, and very low achievement on another. The Student Evaluation Branch has developed procedures for confirming appropriate standards that will allow for the assessment of overall achievement on the test.

The standard for Part A of the Grade 3 English Language Arts test is implicit in the descriptors that were used for scoring (see pages 29-34). The fairness of those descriptors as vehicles for describing reasonable expectations for student writing in the categories of content, development, sentence structure, vocabulary, and conventions was confirmed by the 25 group leaders for the marking session. These group leaders represented each region of the province. Group leaders agreed that writing meeting the descriptions of Satisfactory, Proficient and Exceptional in any of the categories could be deemed to have met or exceeded the minimum expectations of the Grade 3 language arts curriculum. Satisfactory or better scores in each of the five categories comprising Part A would mean a total Part A score of 15/25 (60%) or better. Borderline students, it was agreed, should achieve Satisfactory (3/5) scores in some categories, and Limited (2/5) scores in other categories with equal frequency. Such a combination of Limited and Satisfactory scores should result in a total score of 12/25 (48%), or the minimum standard expected for Grade 3 students.

For Part B of the Grade 3 English Language Arts test, experienced Grade 3 teachers from all parts of the province met to determine what raw score would be expected on the test for a borderline student. The borderline is the division between those who could be expected to achieve the objectives at a minimum level and those who could not. After a review of the curriculum, it was judged that 85% of the Grade 3 students should be able to achieve a minimum level of the objectives of the Grade 3 English Language Arts curriculum, as reflected by the achievement test, given adequate teaching and resources. Since 85% of the students should be able to reach this level, the borderline student would be at the 15th percentile in ability.

The teachers examined each question on the test and determined the difficulty of that question for a 15th percentile student. From the difficulties of individual questions, the overall test difficulty for the borderline student was determined. The average of the test difficulties established by the teachers is the standard for the test. For the Grade 3 English Language Arts Achievement Test, the standard established was as follows: Given the nature and difficulty of this test, 85% of the students should achieve a score of 18/36 or better.

The standard-setting procedures for Part B of the Grade 3 English Language Arts Test were carried out with a provincially representative group of 20 teacher markers during the scoring session for Part A.

Chapter 4

SCORING OF PART A: EXPRESSIVE LANGUAGE - WRITING

Organization of Markers

One hundred and twenty-six teachers from across the province scored Part A: Expressive Language - Writing from July 22 to July 26, 1985 at the Student Evaluation Branch Marking Centre in Edmonton. To qualify for marking, each teacher was required to have a valid permanent Alberta teaching certificate, to have taught Grade 3 English Language Arts for at least two years, and to be currently teaching Grade 3 English Language Arts. In addition, markers were required to have been recommended by their superintendents.

Twenty-five teachers representing all regions of the province were appointed group leaders. They met with Student Evaluation Branch personnel on Friday, July 19, 1985 to prepare for the marking session. This one-day session for group leaders consisted of reading, scoring, and discussing papers that were representative of the range of student writing apparent in the actual samples. The principal focus of the group leaders' discussion was the appropriateness of the scoring guides and their application to the students' writing. In essence, the group leaders validated the standard for assessing Grade 3 students' writing.

Training

The 126 teacher-markers met at the marking centre Monday, July 22. The first morning was used for training. Markers reviewed the scoring guides and marking procedures under the supervision of Student Evaluation Branch personnel. As well, with the guidance of the group leaders and Student Evaluation Branch personnel, markers discussed those papers previously selected (and confirmed by group leaders) to be representative of the range of performance within the sample, and also reflective of the descriptors in the scoring guides.

Scoring

The remainder of the week was used for the scoring of student papers. Each paper was scored independently by one marker. The one-marker system produces results that are reliable at the school and jurisdiction level but not necessarily at the individual student level.

Before the papers were distributed to the markers, student identification was removed and the papers were organized into bundles of 10. Each marker collected a bundle of papers and entered his or her ID number on the back of each paper. The papers were then read and scored independently. Scored papers were rebundled and returned to the clerks who checked score sheets for correct completion. Score sheets were then removed and processed for statistical analysis and reporting.

Although the papers were scored on a one-marker system, 495 papers were recirculated so that for these papers a second set of scores would be available to confirm scoring consistency. Of the scores awarded to the 495 papers on a second reading, 85.9% remained identical to the original score, or varied by only one point, on each of the five scoring scales.

Reliability Reviews

Reliability of results was of prime concern during the scoring sessions, and because of this, reliability review sessions were scheduled twice daily at 10:00 a.m. and at 2:00 p.m. At these sessions each marker was given a copy of the same paper to score independently using the scoring guides. Group discussion of the scores assigned by each marker followed. After the discussion, each marker was given the opportunity to enter a second score in each category. The group leaders forwarded the sets of scores to Student Evaluation personnel who tallied the pre- and post-discussion scores, then posted the resultant distribution of scores. This information provided useful feedback for the markers and helped to ensure greater consistency in the application of the scoring guides.

Group membership was changed at regular intervals during the reliability review sessions.

Chapter 5
RESULTS AND OBSERVATIONS

Test Results

The results of the total test are reported first, followed by separate reports for Part A and Part B.

Results for the Grade 3 English Language Arts Achievement Test

Table 4 shows the frequency distribution of scores for the total test.

Table 4
Grade 3 English Language Arts Achievement Test
Frequency Distribution of Total Scores

Score	Relative Frequency in %*	Cumulative Frequency in %**	Score	Relative Frequency in %*	Cumulative Frequency in %**	Score	Relative Frequency in %*	Cumulative Frequency in %**
0	-	-	34	0.2	2.1	68	3.4	51.9
1	-	-	35	0.3	2.5	69	2.7	54.6
2	-	-	36	0.3	2.8	70	3.7	58.2
3	-	-	37	0.4	3.2	71	2.1	60.4
4	-	-	38	0.3	3.5	72	3.7	64.1
5	-	-	39	0.5	4.0	73	2.3	66.4
6	-	-	40	0.4	4.4	74	3.6	70.0
7	-	-	41	0.7	5.0	75	1.8	71.8
8	-	-	42	0.5	5.6	76	3.3	75.1
9	-	-	43	0.7	6.3	77	1.7	76.8
10	-	0.1	44	0.7	6.9	78	2.8	79.6
11	-	0.1	45	0.8	7.7	79	1.5	81.1
12	-	0.1	46	1.0	8.7	80	2.6	83.7
13	-	0.1	47	0.9	9.6	81	1.2	85.0
14	-	0.1	48	1.1	10.6	82	2.0	86.9
15	-	0.2	49	1.0	11.6	83	1.3	88.2
16	-	0.2	50	1.2	12.9	84	1.6	89.8
17	-	0.2	51	1.2	14.1	85	0.9	90.7
18	-	0.2	52	1.4	15.5	86	1.6	92.3
19	-	0.3	53	1.4	17.0	87	0.7	93.0
20	-	0.3	54	1.5	18.4	88	1.1	94.1
21	0.1	0.4	55	1.8	20.2	89	0.8	94.9
22	-	0.4	56	1.7	21.9	90	1.1	96.0
23	0.1	0.5	57	1.9	23.8	91	0.5	96.4
24	0.1	0.6	58	1.8	25.6	92	0.8	97.2
25	0.2	0.7	59	2.5	28.1	93	0.6	97.8
26	-	0.8	60	1.9	30.1	94	0.7	98.6
27	0.2	0.9	61	2.6	32.6	95	0.3	98.9
28	0.1	1.1	62	2.4	35.0	96	0.5	99.3
29	0.2	1.2	63	2.6	37.7	97	0.4	99.7
30	0.1	1.3	64	2.6	40.3	98	-	99.7
31	0.2	1.6	65	2.7	43.0	99	0.2	99.9
32	0.1	1.7	66	3.1	46.1	100	0.1	100.0
33	0.2	2.0	67	2.4	48.5			

*Relative Frequency: the percentage of students achieving each score.
**Cumulative Frequency: the percentage of students achieving at, or below, each score.

The total test score was obtained by combining the scores for Part A: Expressive Language - Writing and Part B: Receptive Language - Reading so that each part had a weighting of 50%.

Any score that was achieved by fewer than 0.05% of the population is represented by a dash (-). It should be noted, therefore, that the range of student scores was from 0 to 100, although the relative frequency at the lower end of the distribution does not appear to indicate this. Six students achieved a raw score of 0 but, since this represents fewer than 0.05% of the population, the relative frequency is shown as 0.0.

The standard set for minimum level of achievement of the objectives of the Grade 3 English Language Arts Achievement Test was obtained by combining the standard set for Part A and the standard set for Part B. The standard set for the test as a whole was 49%. Eighty-nine per cent of the students achieved this standard.

Results for Part A: Expressive Language - Writing

The average total score on Part A was 62.7%. However, since each category is descriptive of a particular quality of writing, it is most useful to consider scores on a category by category basis.

Table 5 shows the percentage distribution of scores by reporting category for Part A.

Table 5

Part A: Expressive Language - Writing
Percentage Distribution of Scores by Reporting Category

Score	Reporting Category				
	Content	Development	Sentence Structure	Vocabulary	Conventions
5 (Exceptional)	10.3	8.9	7.5	7.5	11.0
4 (Proficient)	25.1	23.3	20.9	20.3	24.9
3 (Satisfactory)	42.1	39.1	49.3	51.7	42.4
2 (Limited)	18.5	24.3	19.0	17.5	17.8
1 (Poor)	3.6	4.0	2.8	2.6	3.5
Ins (Insufficient)	0.4	0.4	0.4	0.4	0.4

Observations regarding the results for Part A are as follows: for CONTENT, 77.5% of the students scored at a Satisfactory level or better, and 22.1% of the students scored at a Limited or Poor level; for DEVELOPMENT, 71.3% of the students scored Satisfactory or better, and 28.3% scored Limited or Poor; for SENTENCE STRUCTURE, 77.7% of the students scored Satisfactory or better, and 21.8% scored Limited or Poor; for VOCABULARY, 79.5% of the students scored Satisfactory or better, and 20.1% scored Limited or Poor; for CONVENTIONS, 78.3% of the students scored Satisfactory or better, and 21.3% scored Limited or Poor. Only 0.4% of the students produced written work that was considered to be Insufficient for scoring purposes. Teacher-markers were generally pleased with the quality of the students' writing. Students handled the narrative form very well, and made their stories interesting through precise vocabulary and creative selection of details. The students often created solutions to Sandy's dilemma in the story-starter that were plausible and yet inventive. Many showed evidence of editing. Such efforts were regarded with approval by the markers.

In summary, overall achievement on the writing assignment was considered acceptable, with at least 71.3% of the students scoring Satisfactory or better on any one marking scale. The highest achievement was on VOCABULARY and the lowest was on DEVELOPMENT.

Table 6 shows the frequency distribution of total scores on Part A: Expressive Language - Writing.

Table 6

Part A: Expressive Language - Writing
Frequency Distribution of Scores

Score	Relative Frequency in %*	Cumulative Frequency in %**	Score	Relative Frequency in %*	Cumulative Frequency in %**
0	0.4	0.4	15	12.9	52.0
5†	1.0	1.4	16	9.0	61.0
6	0.5	1.9	17	8.7	69.7
7	0.7	2.6	18	7.1	76.8
8	1.1	3.7	19	5.5	82.3
9	1.8	5.5	20	4.7	87.0
10	3.8	9.3	21	2.8	89.8
11	4.3	13.6	22	2.7	92.5
12	6.7	20.2	23	2.5	94.9
13	9.0	29.2	24	2.0	97.0
14	9.9	39.1	25	3.0	100.0

The minimum standard on Part A of the Grade 3 English Language Arts Achievement Test was met or exceeded by 86.4% of the students. An explanation of this standard is given on page 8.

---

*Relative Frequency: the percentage of students achieving each score.
**Cumulative Frequency: the percentage of students achieving at, or below, each score.

† Because each of the written response scoring categories is scored from 1 to 5 (POOR to EXCEPTIONAL) it is impossible for a student to have a total raw score of 1 to 4. There were, however, 136 (0.4%) students who wrote nothing in the test booklet, or whose writing was assessed insufficient to score for CONTENT. These were therefore given a score of 0 in each of the five categories.

Results for Part B: Receptive Language - Reading

The average score on Part B was 71.0%.

Tables 7-11 give details about scores according to reporting category, cognitive level, frequency distribution, performance on MACOSA tests, and frequency of response question by question.

Table 7

Results for Part B: Receptive Language - Reading

By Reporting Category

Reporting Category	Number of Questions	Raw Score Mean	Standard Deviation
Total Test	36	25.5	5.9
1. Main Idea	7	4.7	1.4
2. Supporting Detail	7	5.0	1.5
3. Vocabulary	6	4.4	1.4
4. Predictions	8	6.0	1.6
5. Conclusions	8	5.5	1.7

Observations regarding the results for Part B are as follows: for Main Idea, the average score was 4.7/7 (67.1%); for Supporting Detail, the average score was 5.0/7 (71.4%); for Vocabulary, the average score was 4.4/6 (73.3%); for Predictions, the average score was 6.0/8 (75.0%); and for Conclusions, the average score was 5.5/8 (68.8%).

Although performance in the different reporting categories appears to show some variation, caution is advised in comparing them. The sets of questions that make up each reporting category were not selected to be equal in average level of difficulty, therefore differences may be due to variations in question difficulty rather than in student performance. The averages can be used, however, in combination with jurisdictional results to detect patterns of relative strength or weakness in achievement in each of the reporting categories.

Table 8 shows raw scores for Part B by cognitive level.

Table 8

Results for Part B: Receptive Language - Reading

By Cognitive Level

Cognitive Level	Number of Questions	Raw Score Mean	Standard Deviation
Total Test	36	25.5	5.9
Literal Understanding	10	7.1	2.1
Inferential Understanding and Evaluation	26	18.4	4.3

Observations regarding the results for Part B by cognitive level are as follows: for Literal Understanding, students scored an average of 7.1/10 (71.0%); and for Inferential Understanding and Evaluation, students scored an average of 18.4/26 (70.8%).

Because questions within each cognitive level vary in difficulty, and because the average difficulty of the questions in one cognitive level is not necessarily the same as the average difficulty of questions in another cognitive level, no conclusions can be drawn about students' performance on one cognitive level compared to performance on another.

Table 9 shows the frequency distribution of scores on Part B.

Table 9

Part B: Receptive Language - Reading
Frequency Distribution of Scores

Score	Relative Frequency in %*	Cumulative Frequency in %**	Score	Relative Frequency in %*	Cumulative Frequency in %**
0	-	-	19	2.7	15.8
1	-	0.1	20	2.9	18.7
2	-	0.1	21	3.8	22.5
3	-	0.1	22	4.2	26.6
4	-	0.1	23	4.7	31.3
5	0.1	0.2	24	5.2	36.5
6	0.1	0.3	25	6.0	42.5
7	0.2	0.5	26	6.7	49.2
8	0.3	0.8	27	6.8	56.0
9	0.5	1.3	28	7.4	63.4
10	0.5	1.8	29	7.7	71.2
11	0.7	2.6	30	7.7	78.9
12	0.9	3.4	31	6.8	85.7
13	1.0	4.4	32	5.9	91.6
14	1.3	5.7	33	4.2	95.8
15	1.5	7.2	34	2.5	98.3
16	1.7	8.9	35	1.3	99.6
17	1.9	10.8	36	0.4	100.0
18	2.3	13.1			

Any score that was achieved by fewer than 0.05% of the population is represented by a dash (-). It should be noted, therefore, that the range of student scores was from 0 to 36, although the relative frequency at the lower end of the distribution does not appear to indicate this. Fourteen students achieved a raw score of 0 but, since this represents fewer than 0.05% of the population, the relative frequency is shown as 0.0.

The standard set for minimum achievement of the objectives of the Grade 3 English Language Arts Achievement Test (Reading) was met by 89.2% of the students. The procedure used for establishing this standard is described in detail on pages 7 and 8.

---

*Relative Frequency: the percentage of students achieving each score.
**Cumulative Frequency: the percentage of students achieving at, or below, each score.

Table 10 shows the percentages of students choosing correct answers for six MACOSA questions in 1978 and in 1985.

Table 10

Comparison of Performance on MACOSA* Questions
Part B: Receptive Language - Reading

Question Number	1978 Percentage of Students Choosing Correct Answers	1985 Percentage of Students Choosing Correct Answers
31.	75.4	78.4
32.	50.2	44.7
33.	75.0	84.0
34.	68.5	78.7
35.	68.1	74.9
36.	80.8	86.9
	Average = 69.7	Average = 74.7
	Number of Students = 1 408	Number of Students = 32 030

The following observations may be made regarding the scores for these six questions: The differences in the percentages of students choosing correct answers on the six MACOSA questions are statistically significant at the .001 level. A greater percentage of Grade 3 students chose correct answers in 1985 than did Grade 3 students in 1978.

---

*Minister's Advisory Committee on Student Achievement Test (1978).

Table 11 shows question response frequencies for all 36 questions appearing in Part B.

Table 11

Part B: Receptive Language - Reading

Question Response Frequencies

Question Number	Key	Distribution of Responses in %*				Question Number	Key	Distribution of Responses in %*			
		A	B	C	D			A	B	C	D
1.	A	63.4	10.5	12.1	13.8	19.	B	9.9	52.1	18.1	18.4
2.	C	4.5	16.3	75.0	3.9	20.	D	8.1	7.8	11.5	71.4
3.	A	63.5	14.5	8.2	13.3	21.	D	12.9	5.1	4.7	76.1
4.	C	9.6	7.2	72.5	10.3	22.	D	18.4	14.7	9.3	56.0
5.	C	4.6	9.9	80.3	4.5	23.	B	37.8	39.7	15.6	5.4
6.	A	75.6	10.7	8.8	4.6	24.	A	72.0	7.7	7.3	10.9
7.	D	9.8	2.9	2.8	84.0	25.	D	23.1	1.7	3.3	68.8
8.	D	4.2	8.1	1.3	85.8	26.	C	6.5	3.5	85.7	3.1
9.	D	2.2	5.6	24.5	67.1	27.	D	7.9	1.4	5.4	83.7
10.	B	10.7	75.6	7.4	5.7	28.	A	74.4	10.0	7.2	6.9
11.	C	3.0	13.1	78.3	4.9	29.	B	33.6	35.0	15.3	14.4
12.	C	2.4	2.3	93.5	0.9	30.	B	11.1	27.1	12.5	46.9
13.	D	7.8	4.1	11.0	75.7	31.	B	9.5	78.4	7.8	2.2
14.	D	3.3	8.4	2.2	84.9	32.	C	18.9	20.8	44.7	12.9
15.	A	76.8	15.2	5.6	1.7	33.	C	5.9	5.0	84.0	2.5
16.	A	86.5	4.4	6.3	2.1	34.	B	8.6	78.7	7.1	3.6
17.	C	25.2	12.4	54.9	6.5	35.	A	74.9	7.4	5.1	10.1
18.	C	6.2	13.6	69.8	9.3	36.	A	86.9	3.5	3.2	4.1

*The sum of the percentages may not be 100 because the numbers have been rounded.

Discussion of Selected Questions

The results for each reporting category are discussed in detail in the following sections. Those skills that were tested in each category are identified and the easiest and most difficult questions within each category are noted. Sample questions from the test are provided. For each sample question the asterisk(*) indicates the correct response, and the percentage of students who selected each alternative is given.

Main Idea (Questions 5, 9, 14, 19, 30, 31, and 34)

Questions related to this reporting category measure a student's ability to:

- synthesize the main idea of a reading selection
- infer the main idea of a reading selection by using contextual clues
- evaluate the appropriateness of suggested main ideas for a reading selection

The average raw score for the 7 questions on Main Idea was 4.7.

Question 14, requiring students to infer the main idea of a report by using contextual clues, was found to be the easiest (84.9% answered correctly).

Question 30, requiring students to infer the main idea of a story by using contextual clues, was found to be the most difficult (27.1% answered correctly).

A discussion of question 19, a question of greater than average difficulty, follows.

19. Which of the following titles BEST fits the ideas in the story?		Student Responses
O	The Traveller	9.9%
*O	The Secret Map	52.1%
O	The Sad Potter	18.1%
O	The Arabian War	18.4%

Students who selected incorrect answers may not have been able to connect or synthesize several key details from the story.

Supporting Detail (Questions 2, 6, 16, 22, 27, 32, and 35)

Questions related to this reporting category measure ability to:

- recall supporting detail found in a reading selection
- infer the implication of a supporting detail found in a reading selection
- evaluate supporting detail in terms of the main idea

The average raw score for the 7 questions on Supporting Detail was 5.0.

Question 16, requiring students to recall a supporting detail found in a story, was found to be the easiest (86.5% answered correctly).

Question 32, requiring students to recall a supporting detail, found in a short paragraph, was found to be the most difficult (44.7% answered correctly).

A discussion of question 6, a relatively easy question, follows.

6. When did Penny decide to go home?		Student Responses
*O	When the can was full	75.6%
O	When she was in danger	10.7%
O	When the bear chased her	8.8%
O	When she was hot and thirsty	4.6%

This question was answered easily by students, possibly because of the proximity in the text of the description of the can of berries to the statement that Penny was on her way home. Students who answered this question correctly had to be aware of time sequence.

Vocabulary (Questions 1, 10, 15, 20, 21, and 28)

Questions related to this reporting category measure ability to:

- recall the meanings of words and expressions
- infer word meaning from context
- evaluate the appropriateness of word usage

The average raw score for the 6 questions on Vocabulary was 4.4.

Question 15, requiring students to infer a word meaning from context, was found to be the easiest (76.8% answered correctly).

Question 1, requiring students to recall the meaning of a word, was found to be the most difficult (63.4% answered correctly).

A discussion of question 20, a question of average difficulty, follows.

20. What does the underlined word <u>dainty</u> mean?		Student Responses
O	Deep	8.1%
O	Round	7.8%
O	Heavy	11.5%
*O	Gentle	71.4%

Although there are indirect contextual clues such as "silky, "satiny" and "elegant" pointing toward the right answer, this question was one which required the student to rely upon background knowledge. Students had to recall the meaning of the word "dainty" from all of their experiences in living, and many of them were able to do so.

Predictions (Questions 4, 8, 12, 18, 24, 29, 33 and 36)

Questions related to this reporting category measure ability to:

- determine, through inference, the outcome of a series of events
- evaluate a suggested possible outcome of a series of events

The average raw score for the 8 questions on Predictions was 6.0.

Question 12, requiring students to evaluate the suggested possible outcome of a series of events, was found to be the easiest (93.5% answered correctly).

Question 29, also requiring students to evaluate the suggested possible outcome of a series of events, was found to be the most difficult (35.0% answered correctly).

A discussion of question 29 follows.

29. What did the elf MOST LIKELY do after the shoemaker went into his storeroom?		Student Responses
O	He made another pair of shoes.	33.6%
*O	He left the shop to help someone else.	35.0%
O	He took all the shoes from the window.	15.3%
O	He brought the friends of the shoemaker into the shop.	14.4%

As with all evaluation questions, the four alternatives in question 29 have a measure of correctness or plausibility. Such questions require the students to select the MOST LIKELY alternative. Students who incorrectly chose alternative A may have done so because of its position (first) and its plausibility. It is likely that those students did not re-read any parts of the story in order to confirm their answers. The second last paragraph strongly suggests that the shoemaker was not deserving of the elf's help because he was greedy.

Conclusions (Questions 3, 7, 11, 13, 17, 23, 25, and 26)

Questions related to this reporting category measure ability to:

- draw (infer) appropriate conclusions from details and ideas present in a reading selection
- evaluate the relative importance of suggested concluding statements

The average raw score for the 8 questions on Conclusions was 5.5.

Question 26, requiring students to infer an appropriate conclusion, was found to be the easiest (85.7% answered correctly).

Question 23, also requiring students to infer an appropriate conclusion, was found to be the most difficult (39.7% answered correctly).

A discussion of question 23 follows.

23.	Why does the poet call the cat "Beauty"?	Student Responses
O	The poet likes the cat because it is pretty.	37.8%
*O	The poet likes the cat even though it is ugly.	39.7%
O	The poet likes the cat because she wants a pet.	15.6%
O	The poet likes the cat because it is Siamese.	5.4%

Question 23 directs students to select an answer that is a correct conclusion. Students who incorrectly selected alternative A may have done so because of its position (first), and/or because of the association of "Beauty" with prettiness. Re-reading would have helped students to confirm their choice of answer; the cat is described as a "splotchy-blotchy/city cat, not a pretty cat...". Students may have missed the word "not" in the line "not a pretty cat."

Summary of Observations

The provincial average for the Grade 3 English Language Arts Achievement Test was 66.9%. Eighty-nine per cent of the students met or exceeded the minimum standard of 49.0%.

At the provincial level, student achievement on Part A: Expressive Language - Writing was acceptable with an average total score of 62.7%. Eighty-six per cent of the students met or exceeded the minimum standard. More than 77% of the students scored at the Satisfactory level or better in Content, Sentence Structure, Vocabulary and Conventions (77.5%, 77.7%, 79.5%, and 78.3%). Just over 71% of the students scored at the Satisfactory level or better in Development (71.3%).

The provincial average for the 36 multiple-choice questions on Part B: Receptive Language - Reading was 71.0%. The expected performance for Part B was that 85% of the students would score 50.0% or higher. Student performance exceeded that expectation since 89.2% of the students scored 50.0% or higher.

On the six MACOSA questions, a greater percentage of students chose correct answers in 1985 than did Grade 3 students in 1978. The average score on these questions was 74.7% as compared to the average 1978 score of 69.7%.

Chapter 6

GUIDE TO THE INTERPRETATION OF JURISDICTION RESULTS

In addition to their use in monitoring student achievement for the province as a whole, the results of the Grade 3 English Language Arts Achievement Test are useful in comparing achievement in a particular jurisdiction with provincial results. However, care must be exercised in making these comparisons and in drawing conclusions from the data.

The following jurisdiction and school reports are provided for each jurisdiction under separate cover.

1. The Jurisdiction Summary Report contains jurisdiction equivalents of the provincial results that are given in all statistical tables in this report.

2. The School Summary Report contains the school equivalents of the provincial results that are given in all statistical tables in this report.

3. Individual Student Subtest Results are reported for each school.

These reports are confidential to the jurisdiction.

Differences Between Jurisdiction and Provincial Averages

Jurisdictions are provided with their average scores for each reporting category. These scores may be compared to the provincial average for the same reporting category. However, the importance of differences between group jurisdiction averages and provincial averages is not always clear. To aid in the interpretation of differences between the averages, jurisdiction and school reports indicate when the difference is unlikely to be due to chance variation in the abilities of students. For the purposes of the provincial testing program, the 95% confidence interval is used. That is to say, if the probability is less than 1 in 20 that the difference is due to chance, the difference is very likely a real difference, and the jurisdiction average is classified as different from the provincial average. Otherwise, it is classified as not different from the provincial average. The provincial average for that reporting category determines the true population average. The standard deviation for the jurisdiction is used to estimate the standard error of the mean.

Because achievement levels are calculated by taking jurisdiction size into consideration, two jurisdictions with the same averages but of different sizes may be classified differently. The larger jurisdiction would be more likely to be above or below average, because the amount of chance variation would be less in larger jurisdictions, and the actual difference would represent a larger variation from the provincial average.

As an example, imagine two jurisdictions: A has 25 students writing Test X; and B has 100 students writing Test X. Both jurisdictions have the same average, 54.2; both jurisdictions have a standard deviation of 12.0. Test X has a provincial average score of 50.0. The difference between the provincial average and the jurisdiction average is 4.2. A difference this large would be expected 8 times out of 100 for groups of 25 selected at random from the population, and fewer than 3 times out of 1000 for groups of 100. Thus the difference from the provincial average would not be statistically significant for Jurisdiction A, but would be for Jurisdiction B.

For the criterion-referenced scales, which are reported as frequency distributions, a slightly different test of significance is used. The proportion of students scoring satisfactory or better (3 or higher) in the jurisdiction is compared with the same proportion for the province, and a chi square test of significance is performed to determine whether the differences could be due to chance. Again, the 0.05 level of significance is used.

When it has been determined that a difference is significant, the direction of the difference is important, particularly for those jurisdictions below the provincial average. These jurisdictions are encouraged to identify the sources of these differences.

School reports contain the same analysis to determine whether the school varies significantly from the provincial mean.

Table 12 on page 27 indicates the percentage of jurisdictions classified as significantly above or below the provincial average for each reporting category.

Table 12
Distribution of Jurisdiction Levels of
Achievement on the Grade 3 English Language Arts Achievement Test

Reporting Category	% Below the Provincial Average	% Not Different From Provincial Average	% Above the Provincial Average
Total Test	15.0	65.3	19.7
Total Written Response	16.9	66.7	16.4
Total Reading	11.3	63.7	24.5
Main Idea	10.5	76.6	12.9
Supporting Detail	8.2	74.5	17.3
Vocabulary	14.0	72.0	14.0
Predictions	11.9	69.0	19.0
Conclusions	11.0	70.0	19.0
Literal	9.6	68.4	22.0
Inferential and Evaluation	12.7	65.1	22.2

A test score does not indicate why a particular performance occurred, only that it did occur. Identification of reasons for that performance should be undertaken once results have been studied. There are a variety of factors that should be examined:

1. Student motivation. Consideration should be given to the degree to which students were motivated to perform to their levels of ability.

2. Student ability. While the notion of a target region is designed to take into consideration year-to-year fluctuations in the average ability levels of students, it is possible that a group of students with a particularly high or low average ability may come through a system. This is much more likely to be a factor in small systems than in large ones.

3. Teaching and curriculum. Consideration should be given to the type of instruction students have received in the jurisdiction and the adequacy of curricular implementation.

In examining the test results, the reader must keep in mind that a test score does not indicate why a particular performance occurred, but only that it did occur. After studying the results, the identification of reasons for that performance should be undertaken. There are a variety of factors that should be examined:

1. Student motivation. Consideration should be given to the degree to which students were motivated to perform to their levels of ability.

2. Student ability. While the statistical test of significance is designed to take into consideration fluctuations in the average ability levels of students, it is possible that a group of students with a particularly high or low average ability may come through a system. This is much more likely to be a factor in small systems than in a large one.

3. Teaching and curriculum. Consideration should be given to the type of instruction students have received in the jurisdiction and the adequacy of curricular implementation.

There will be other factors that are of importance in particular jurisdictions. School boards wishing to examine further the results in light of local factors are encouraged to establish their own local interpretation panels.

Absentee Rates

If more than 10% of the eligible students in a jurisdiction did not write the test, the reported averages for that jurisdiction may not accurately represent the true averages. Teacher-assigned marks for students who did not write could be compared with teacher-assigned marks for students who did write. If the averages are the same for the two groups, the reported achievement averages are probably representative. If the averages are different, some estimates can be made of what the achievement averages might have been if all students had written the test. Jurisdictions with high absentee rates may wish to contact the Student Evaluation Branch for assistance in estimating their averages.

APPENDIX A

GRADE 3 SCORING GUIDES: EXPRESSIVE LANGUAGE - WRITING

Scoring Guides: Expressive Language - Writing (Definitions and Discussion)

CONTENT	That which the student chooses to write about. This includes the WHO, WHAT, WHERE, and WHEN of a story.
DEVELOPMENT	The method by which the student chooses to organize the content. This includes the sequence of events by which the student organizes the story into a coherent whole. The sequence may involve ordering by cause and effect, but more usually involves a sequence ordered by time and/or by place.
SENTENCE STRUCTURE	The forms of the sentences that the student uses. The category "sentence structure" includes the types of sentences, co-ordination (i.e. linkage of clauses, e.g., "...and so..." or "...but..."), subordination (e.g., "...because..." or "...when..."), and the arrangement within a sentence (e.g., subject/verb/object). Sentence fragments and/or run-on sentences are considered to be part of "sentence structure" rather than "conventions."
VOCABULARY	The words chosen by the student. The vocabulary category considers the precision and clarity of word choice (e.g., "mumbled" instead of "said," "canoe" instead of "boat," "wicked" instead of "bad").
CONVENTIONS	The way in which the student uses standard conventions of language. This includes the use of standard spelling, punctuation, and capitalization.

Since there are five categories in which a student's writing will be evaluated, it is not expected that a student's writing will be equally strong in each category. It is entirely possible that a piece of writing might have been scored as PROFICIENT in one of the five categories but might have been scored as SATISFACTORY or LIMITED in one of the other four categories.

Reporting Category: CONTENT*

(Selecting Details Appropriate to Purpose)**

SCORE		DESCRIPTION OF PERFORMANCE
5	EXCEPTIONAL	A context is clearly established, and events are plausible within that context. Events and actions are sometimes connected implicitly to character motivation. Characters act consistently. Many precise and appropriate details establish characters and events even though experiences may be of an everyday nature.
4	PROFICIENT	A context is clearly established, and most events are plausible within that context. Events and actions are occasionally connected implicitly to character motivation. Characters usually act consistently. Many appropriate details establish characters and events even though experiences may be of an everyday nature.
3	SATISFACTORY	A context is clearly established, and some events are plausible within that context. Events and actions are infrequently connected to character motivation. Characters sometimes act consistently. Some appropriate details establish characters and events even though experiences may be of an everyday nature.
2	LIMITED	A context is vaguely established, and few events are plausible within that context. Events and actions are rarely connected to character motivation. Characters may act inconsistently. Few details establish characters and events.
1	POOR	A context is unclear or missing, but events may be plausible. There is a lack of appropriate detail.
Ins	INSUFFICIENT	Too little writing exists for a judgment to be formed.

*That which the student chooses to write about. This includes the WHO, WHAT, WHERE, and WHEN of a story.
**Details selected by the student will be either descriptive or narrative and associated with characters or events.

Reporting Category: DEVELOPMENT*

(Organizing Details into a Coherent Whole)

SCORE		DESCRIPTION OF PERFORMANCE
5	EXCEPTIONAL	Events have been placed in a coherent and recognizable sequence. The story's unity is strengthened by details about character and actions. Digressive details, if present, do not interfere with the development of the story. Appropriate closure has been achieved.
4	PROFICIENT	Events have been placed in a coherent sequence. The story's unity is sometimes supported by details about characters and actions. Digressive details, if present, do not interfere with the development of the story. Closure has been achieved.
3	SATISFACTORY	Events have been placed in a generally coherent sequence. Digressive details may begin to interfere with the story's development. Closure has been attempted.
2	LIMITED	A sequence of events can be detected, but coherence is not achieved. Digressive details interfere with the unity of the story. Closure, if attempted, may be unsuccessful.
1	POOR	No coherent sequence of events is apparent. Digressive details, if present, interfere greatly with the unity of the story. A sense of closure may be missing, or is incomplete.
Ins	INSUFFICIENT	Too little writing exists for a judgment to be formed. Writing that has been awarded an "Ins" for CONTENT is insufficient.

*The method by which the student chooses to organize the content. This includes the sequence of events by which the student organizes the story into a coherent whole. The sequence may involve ordering by cause and effect, but more usually involves a sequence ordered by time and/or by place.

Reporting Category: SENTENCE STRUCTURE*

(Structuring Sentences Appropriately)

SCORE		DESCRIPTION OF PERFORMANCE
5	EXCEPTIONAL	Appropriate and/or purposeful variation in sentence type, length, and structure is evident. Co-ordination and subordination have been used appropriately and/or sometimes deliberately. Sentence fragments and/or run-on sentences, if present, do not impede meaning.
4	PROFICIENT	Some appropriate and/or purposeful variation in sentence type, length, and structure is evident. Co-ordination and subordination are used appropriately. Co-ordination may be predominant. Sentence fragments and/or run-on sentences, if present, do not impede meaning.
3	SATISFACTORY	Occasional appropriate and/or purposeful variation in sentence type, length, and structure is evident. Co-ordination is used extensively and appropriately. Some subordination may be present. Sentence fragments and/or run-on sentences, if present, do not impede meaning.
2	LIMITED	Little appropriate and/or purposeful variation in sentence type, length, and structure is evident. Co-ordination has been overused, sometimes inappropriately. Sentence fragments and/or run-on sentences, if present, impede meaning.
1	POOR	Co-ordination has been used almost exclusively and inappropriately. Sentence fragments and/or run-on sentences, if present, severely impede meaning.
Ins	INSUFFICIENT	Too little writing exists for a judgment to be formed. Writing that has been awarded an "Ins" for CONTENT is insufficient.

*The forms of the sentences that the student uses. The category, sentence structure, includes the types of sentences, co-ordination (i.e., linkage of clauses, e.g., "...and so..." or "...but..."), subordination (e.g., "...because..." or "...then..."), and the arrangement within a sentence (e.g., subject/verb/object). Sentence fragments and/or run-on sentences are considered to be part of "sentence structure".

Reporting Category: VOCABULARY*

(Selecting and Using Words and Expressions Appropriately)

SCORE		DESCRIPTION OF PERFORMANCE
5	EXCEPTIONAL	Precise and specific verbs, nouns, and/or modifiers have been used appropriately to create clear images.
4	PROFICIENT	Some specific verbs, nouns, and/or modifiers have been used appropriately to create clear images.
3	SATISFACTORY	Words are general but varied and correct. Few specific verbs, nouns, and/or modifiers have been used to create clear images.
2	LIMITED	General verbs, nouns, and/or modifiers have been used correctly. Images are vague.
1	POOR	General verbs, nouns, and/or modifiers have been used, often incorrectly or repetitively. Images are unclear.
Ins	INSUFFICIENT	Too little writing exists for a judgment to be formed. Writing that has been awarded an "Ins" for CONTENT is insufficient.

*The words chosen by the student. The vocabulary category considers the precision and clarity of word choice (e.g., "mumbled" instead of "said," "canoe" instead of "boat," "wicked" instead of "bad").

Reporting Category: CONVENTIONS*

(Following the Conventions of Written Language Appropriately)

SCORE		DESCRIPTION OF PERFORMANCE
5	EXCEPTIONAL	Control of spelling, punctuation, and capitalization facilitates clear communication. Misspellings are easily decipherable. Dialogue, if present, may not be punctuated properly.
4	PROFICIENT	General control of spelling, punctuation, and capitalization facilitates clear communication. Misspellings are decipherable. Dialogue, if present, may not be punctuated properly.
3	SATISFACTORY	Some control of spelling, punctuation, and/or capitalization facilitates communication. Misspellings are generally decipherable. Dialogue, if present, may not be punctuated properly.
2	LIMITED	Lack of control of spelling, punctuation, and/or capitalization generally interferes with communication. Misspellings may be undecipherable. Dialogue, if present, may not be punctuated properly.
1	POOR	Lack of control of spelling, punctuation, and/or capitalization severely interferes with communication. Misspellings are generally undecipherable. Dialogue, if present, may not be punctuated properly.
Ins	INSUFFICIENT	Too little writing exists for a judgment to be formed. Writing that has been awarded an "Ins" for CONTENT is insufficient.

*The way in which the student uses standard conventions of language. This includes the use of standard spelling, punctuation, and capitalization.

APPENDIX B

GRADE 3

SAMPLE SCORE SHEET

Marker ID number: _____

CONTENT	0	1	2	3	4	5
DEVELOPMENT	0	1	2	3	4	5
SENTENCE STRUCTURE	0	1	2	3	4	5
VOCABULARY	0	1	2	3	4	5
CONVENTIONS	0	1	2	3	4	5